Euclidean Distance between original and perturbed embeddings
AdvGLUE attack

Legend:
- Input space - ChatGPT
- Output space - ChatGPT
- Input space - FLAN T5
- Output space - FLAN T5